DOCUMENT RESUME

ED 286 373 FL 016 923

AUTHOR Koopmans, Matthijs
TITLE The Difference between Task Understanding and Reasoning Skills in Children's Syllogistic Performance.
SPONS AGENCY Office of Bilingual Education and Minority Languages Affairs (ED), Washington, DC.
PUB DATE 24 Apr 87
NOTE 27p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, DC, April 24, 1987). Document contains light type.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Abstract Reasoning; Age Differences; *Bilingual Students; Cognitive Processes; Comparative Analysis; Elementary Education; *Elementary School Students; English; Learning Strategies; *Logical Thinking; *Problem Solving; Puerto Ricans; *Reaction Time; Response Style (Tests); Spanish
IDENTIFIERS *Syllogistic Reasoning

ABSTRACT

A study investigated the reason for the high performance of bilingual children on a syllogistic reasoning task. Thirty-nine Puerto Rican Spanish-English bilingual children in grades 3 through 6 were asked to solve problems in both languages. Response times were recorded, and the subjects were asked to justify their answers to assess to what extent reasoning led to task success. Response times were found to be longer if the justification given reflected reasoning, but it was also found that students needed more time to give an answer in their stronger language (Spanish) than in their weaker one (English). In addition, solution accuracy and solution strategy did not interact and the strategies used to solve syllogisms were manifested differently at different grade levels. In the lower grades, the justification given seemed to reflect whether or not the student understood the purpose of the task, and in higher grades, the justification appeared to show whether the student succeeded in his reasoning effort or used a default strategy to solve the problem. Performance above the chance level is attributable to the characteristics of the syllogism rather than to what the subjects did with them, suggesting that strategy rather than success is a better source of information about children's performance. (MSE)

The difference between task understanding and reasoning skills in children's syllogistic performance.

Matthijs Koopmans(1)
Harvard University

Paper presented at the AERA Annual Conference, Washington, DC April 24, 1987.

---

[1] Research reported in this paper is partly funded by the Office for Educational Research and Improvement through the auspices of the Center for Language Education and Research

ABSTRACT

In this study, an attempt has been made to find an explanation for the fact that bilingual elementary school children performed above chance level on a syllogistic reasoning task, regardless of whether they appear to reason or not. In order to assess to what extent reasoning led to task success, subjects were asked to justify their answers. The prediction was made that response times for the answers would be higher if the justification given reflected reasoning. Thirty nine Puerto Rican bilinguals from grade 3 through 6 were tested in English and Spanish. It was found that response times were indeed longer if the justification given reflected reasoning, but it was also found that subjects needed more time to give an answer in their stronger language (Spanish), than in their weaker one (English). In addition, it appeared that solution accuracy and solution strategy did not interact. Moreover, the analysis indicated that the strategies used to solve syllogisms manifest themselves differently for different grade levels. In the earlier grades, the explanation given seemed to reflect whether the subjects understood the purpose of the task, whereas in the higher grades, it appeared to indicate whether subjects succeeded in their reasoning effort, or used a default strategy to solve the problem. Performance above chance level, finally, is likely to be attributable to the characteristics of the syllogisms rather than to what the subjects did with them. It is argued on the basis of these results that strategy use is a better source of information to assess children's performance than task success.

As long as logic has been established as a formal set of rules, there has been an interest in the extent to which individuals possess the skills to apply those rules successfully for problem solving. The acquisition of the capacity to reason according to logical principles is generally considered to be an important step in the cognitive development of children. In Piaget's developmental theory, it is the last step taken toward cognitive maturation (e.g. Piaget, 1967). Logical reasoning enables individuals to perform mental operations on an abstract level, and thus to emancipate themselves from their direct concrete environment.

Performance on syllogistic tasks is often seen as a typical instance of these formal operations, since a correct solution to these tasks can only be reached by inferring a solution from the propositions as they are given in the premises, propositions which, moreover, are often hypothetical (Cole & Scribner, 1981; Orasanu & Scribner, 1982). On the basis of Piaget's theory, one would be led to the prediction that children are never able to solve syllogisms whereas grown ups would always be able to do so. Neither appears to be the case: Luria (1976) found that illiterate grown ups were unable to solve syllogisms whereas Hawkins, Pea, Glick & Scribner (1984) found that under some conditions, 4 and 5 year old children were able to solve them.

Luria (1976), who did a series of studies in the early

thirties to investigate the effects of social and economical changes brought about by the Soviet revolution on the cognitive and perceptual skills of the inhabitants of rural areas, found that many illiterate peasants were unable to solve syllogisms. In addition, he found that performance on syllogisms that were directly related to the personal experience of the subjects was substantially better than when the solution to the problems could only be reached on the basis of inference from the premises. On the other hand, a group of subjects with short term schooling were able to solve both types of syllogisms without any difficulty. Cole, Gay, Glick & Sharp's (1971) findings on the Kpelle, a population of predominantly illiterate rice farmers in Liberia were consistent with Luria's results: the nonliterate groups showed many errors and much misunderstanding in most cases, whereas the schooled subjects did substantially better.

To get a clearer sense of what went wrong, the errors in syllogistic reasoning have been extensively analyzed, and they appear to be attributable to the fact that unschooled subjects tend to solve syllogisms by relying on their personal experience rather than by recognizing the premises in the syllogisms as the key to the right answer. For example, when Luria's subjects were confronted with syllogisms like the following:

> Cotton grows where it is hot and dry.
> In England, it is cold and damp.
> Does the cotton grow in England? (Luria, 1976, p 107)

they would typically answer:

> We don't know that, we know that it grows in our country (ib., p.111).

or:

> It's chilly here too (ib., 1976, p. 111).

Subjects appeared to understand the syllogism task as a test of their knowledge of facts, rather than as a test of their reasoning skills, and consequently, their performance could not be attributed to faulty reasoning, but rather to a misunderstanding of the purpose of the task as the experimenter had it in mind. The importance of Luria's result lies in its methodological implications; it illustrates that a mutual agreement between subject and experimenter about the purpose of the task is necessary in order to obtain an interpretable result.

On the basis of Luria's study, a distinction has been made in subsequent work among the types of justification given for the

answers. These justifications have normally been categorized as theoretical versus empirical explanations (Scribner, 1977). An empirical explanation is an explanation in which subjects rely on their personal experiences to solve the reasoning tasks, whereas in a theoretical explanation, subjects refer to the premises in their justifications. Empirical explanations have generally been taken as an indicator that subjects interpret the task as a test of their knowledge of facts, and theoretical explanations are taken as a reflection of task understanding, and as an indication that subjects use their reasoning skills to solve syllogisms.

Scribner (1977) and Hawkins et al. (1984) indeed found that empirical solutions were usually wrong, and that theoretical solutions were usually correct. My study (Koopmans, in preparation) appears to contradict this finding: subjects performed above chance level regardless of whether they justified their answers theoretically or empirically. Furthermore, it appeared that most subjects did not consistently give theoretical or empirical justifications to their answers. Although it is probably safe to conclude that subjects do not understand the task if none of their answers is theoretical, and equally safe to conclude that they do understand it if all answers are theoretically justified, it is harder to make such inferences on the basis of subjects' explanation type if some of the answers are theoretically justified, and some empirically. The question whether and how task understanding reflects reasoning, thus,

deserved closer scrutiny. The purpose of the present study is to investigate this question.

The present study

Since neither solution accuracy nor explanation type appears to be a sufficient source of information to establish whether reasoning takes place, an additional source of information has been sought. It has been hypothesized that if the distinction between theoretical and empirical justifications reflects reasoning, processing of the syllogisms will be slower for answers that were justified theoretically than for answers which were justified empirically. Slower processing will reflect itself in higher response times. This hypothesis is based on the finding that reasoning is a more complex mental operation than decoding task input and matching it against everyday knowledge (Sternberg, 1986). It has further been hypothesized that if reasoning increases the likelihood of giving the right answer, as Scribner (1977) found, there should be a similar difference between response times for right answers and wrong answers. To account more fully for the effect of task understanding, bilingual subjects have been used in the study to examine whether there is a difference in this respect according to the language in which the syllogisms were administered.

Subjects in this study were 39 Puerto Rican elementary school children who were tested in English as well as Spanish. Ten syllogisms were administered in random order in each language. Third through sixth graders were included in the study. Response times were measured in the following way: the stopwatch was started at the point where the experimenter started the question, and it was halted upon the subjects confirmation or disconfirmation.

The children were told that a number of stories would be read to them with a question which they had to answer as fast as they could, and afterwards explain why they gave this answer. Since no invalid syllogisms were included in the study, subjects had the option of answering 'yes' or 'no'.

Results

Since the primary concern of this study was to assess by the item whether explanation type related to solution accuracy, an analysis of variance has been performed on response times (n=672) which included 2 categories for solution accuracy (right, wrong), 3 categories for explanation type (theoretical, empirical, "don't know"), 2 levels for language of administration (English, Spanish), and 4 levels for grade (3rd, 4th, 5th, 6th).

The analysis revealed significant main effects for explanation type ($F=9.46$, $df=2,669$, $p=.0001$), and for language of administration ($F=56.81$, $df=1,670$, $p=.0001$). It does make a difference in response time, then, whether subjects justified their answers to the syllogisms empirically or theoretically, but it also made a difference whether syllogisms were administered in English or Spanish. The two factors did not interact, however. It took subjects longer to come up with an answer if syllogisms were administered in Spanish rather than English, and it took them longer to give answers that were theoretically justified than answers that were empirically justified. More cognitive effort is invested, then, in the syllogisms that received a theoretical answer, suggesting that subjects do apply reasoning skills in those cases, whereas in empirically justified answers they do not, or to a lesser extent. However, subjects also appear to take more time to give an answer if the syllogisms are administered in Spanish rather than English. It seems that more cognitive effort is devoted to the task if subjects reason in their stronger language. The absence of an interaction between language of administration and solution strategy indicates, however, that the difference in response times according to solution strategy is not affected by language of administration. The extra cognitive effort that is required to give an answer which is theoretically grounded, thus, is not attributable to the language in which the syllogism has been administered. Conversely, the extra time needed to solve the syllogisms in Spanish cannot be explained in

terms of the type of justification given for the answer.

No significant main effects were found for grade and solution accuracy, indicating that response times neither vary according to the grade level of the subjects, nor to the correctness of their answer. Particularly this latter finding is striking since it suggests that whether reasoning takes place or not is not affected by the accuracy of the solution obtained. In other words, solution accuracy does not appear to reflect whether subjects use reasoning to meet the task demands.

The analysis reveals a significant interaction, however, between solution accuracy and language of administration ($F=3.30$, $df=1,670$, $p=.0687$ (see figure 1). In English, there is only a slight difference in response times for the wrong and the right answers, whereas in Spanish, subjects took much longer for the wrong answers. In their stronger language, subjects appear to try harder on the syllogisms for which they end up giving the wrong answer than for those for which their answer is right, whereas no such difference exists in the weaker language (English). In other words, reasoning is reflected in solution accuracy only in Spanish.

More importantly, the analysis revealed a significant interaction between language and grade ($F=5.93$, $df=3,668$, $p=.0005$). This interaction is plotted in figure 2. A reverse

FIGURE 1 Average response times in English and Spanish for correct and incorrect responses

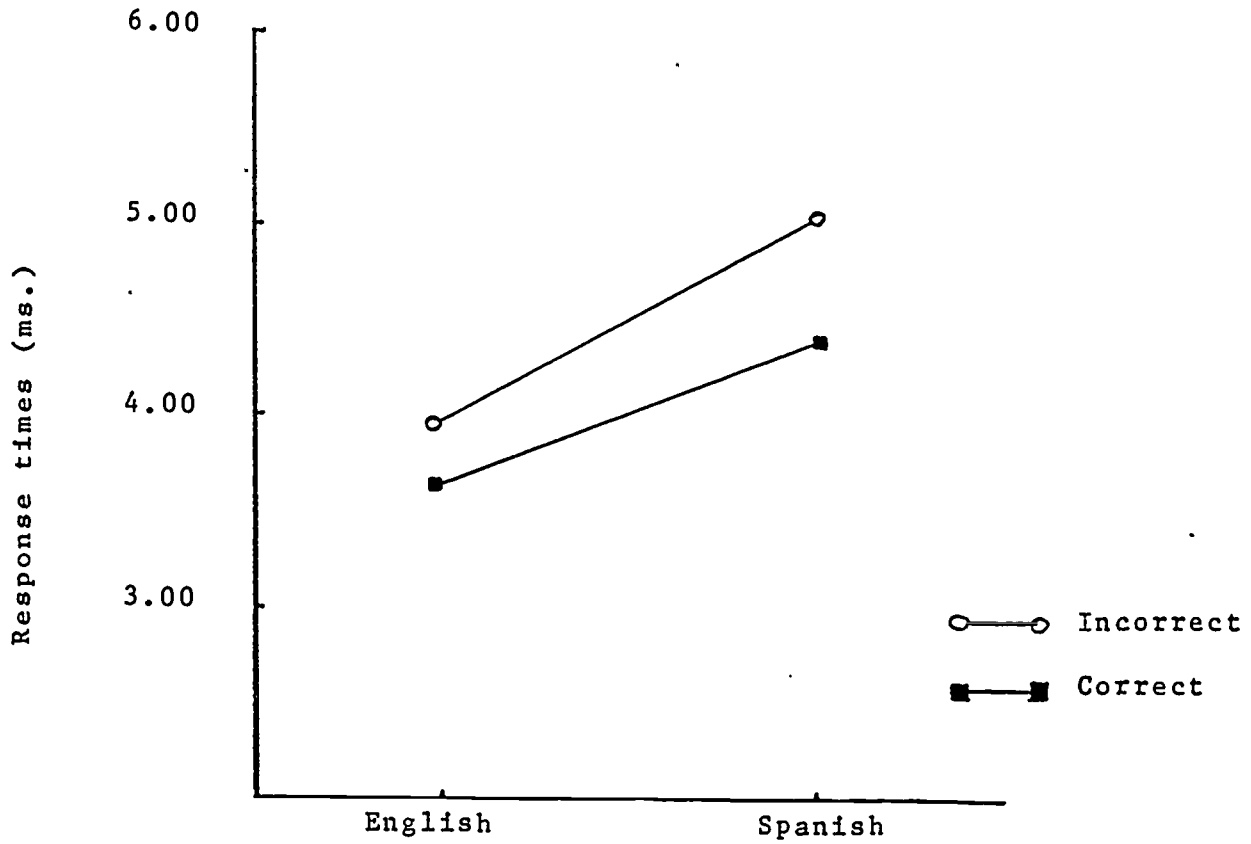
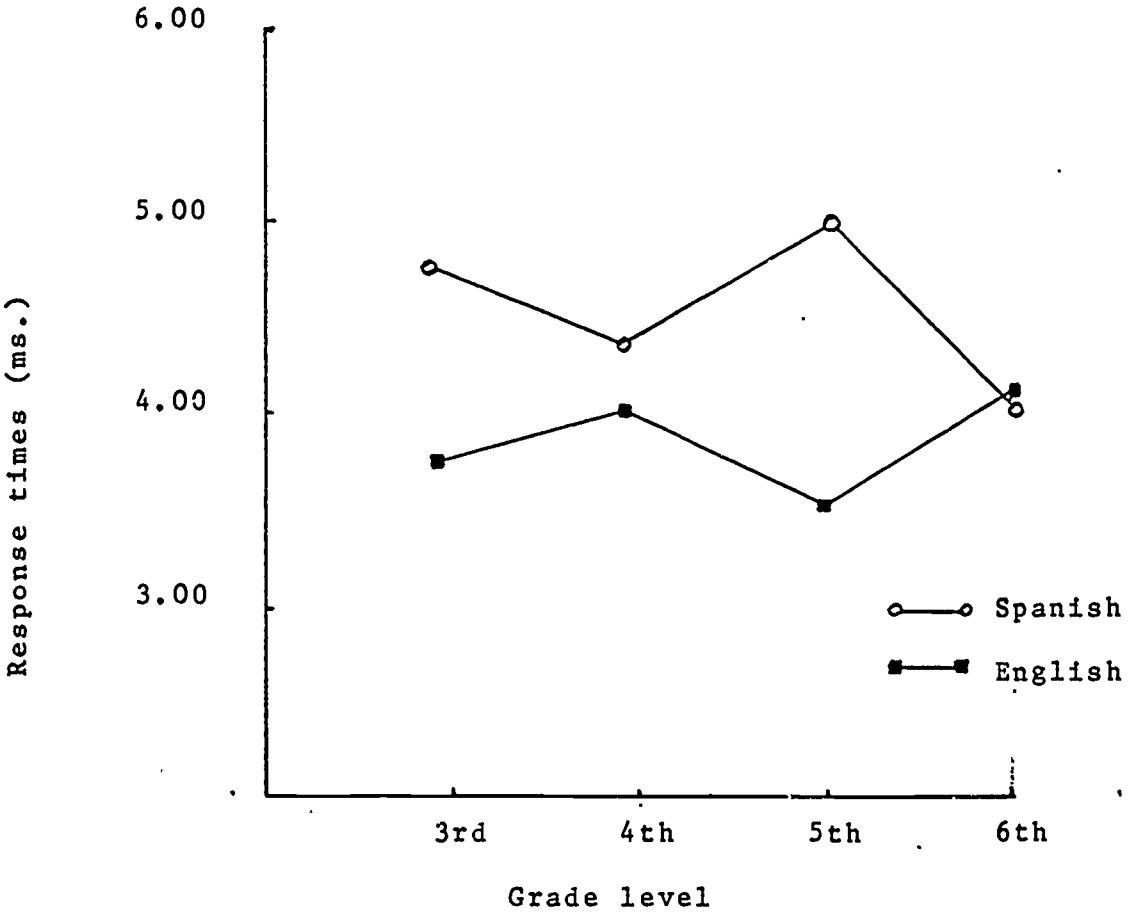


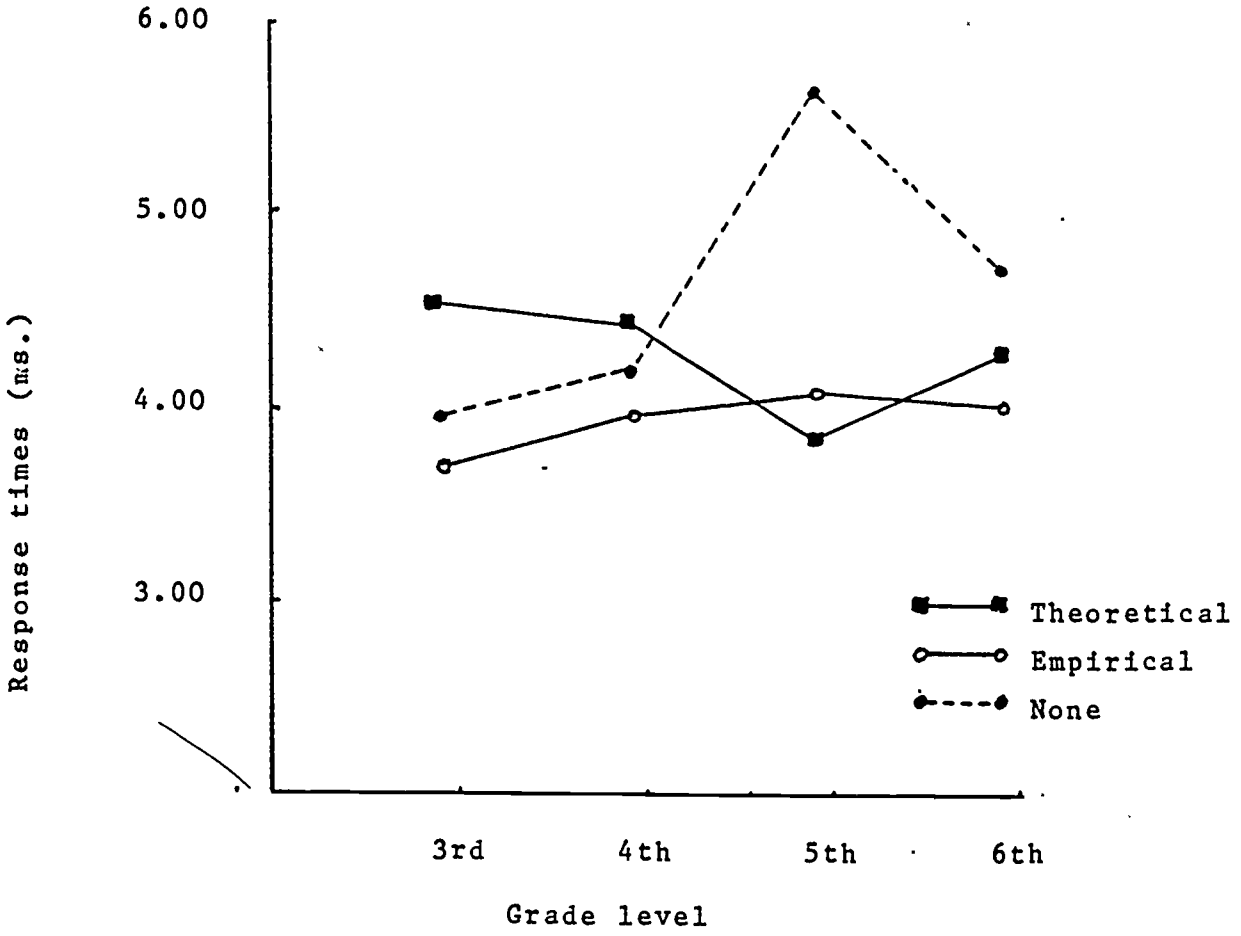
FIGURE 2 Average response times in English and Spanish for each grade



tendency according to grade can be noted in the response times in Spanish and English. Higher response times in English tend to go along with low response times in Spanish, and vice versa: the more time subjects need in English to solve the syllogisms the less time they tend to use in Spanish. Given the significance of the main effect for language of administration, this interaction is probably a reflection of the difference in language proficiency of the children in different grades.

The interaction between explanation type and grade is also significant ($F=2.73$, $df=6,665$, $p=.0127$). It can be seen in figure 3 that response times for theoretical explanations tend to go down as grade level gets higher, whereas for empirically explained answers, the response times tend to stay the same for the different grades. For the 'don't know' category, however, there is an increase in response times according to grade, indicating that much thought is given to those syllogisms for which subjects are unable to explain their answer in the higher grades, but not in the lower ones. This tendency suggests that in the higher grades, most notably the fifth, reasoning does underlie this explanation type, whereas in the lower grades (3rd and 4th), it does not. There may be two reasons why this tendency is not as clear for the sixth grade is it is for the fifth. First, it may be due to random factors, since there were only 4 sixth graders included in this sample. Second, the sixth grade subjects in this sample were relatively weaker in English than

FIGURE 3 Average response times for each grade according to type of justification



children in lower grades, which may have led to a profile for hem which is more comparable to that of the earlier grades than that of the fifth grade. This possibility will be discussed more extensively below.

Most noteworthy is the absence of an significant interaction between solution accuracy and explanation type ($F=.063$, $df=2,669$, $p=.5920$). Apparently, response time for the right answers and for the wrong answers is not differentially affected by explanation type, although explanation type, taken by itself, does make a significant difference: whether the child gives the right or the wrong answer does not depend on whether reasoning takes place, although type of explanation does so regardless of whether it leads to the right answer or not.

Discussion

This study was designed to find an answer to the question why subjects performed above chance level on syllogistic reasoning tasks regardless of whether their explanations suggested they understood the task, and to find a way to assess task understanding of those subjects who did not consistently give theoretical or empirical explanations. The response times of the subjects were used as an independent source of evidence and an analysis of variance was performed on these response times. Before discussing these response times more extensively, it

should be pointed out that this design does, formally speaking, not allow for inferences about the performance of the population of Puerto Rican bilingual elementary school children, since the observations which constituted the sample of this study were response times per item of the syllogism. In other words, for each subject, there were up to 20 observations, and the population to which inferences can be made is a population of response times. This procedure has been followed because the majority of subjects, 27 out of 39, did not consistently give theoretical or empirical explanations, and taking mean response times for each subject and an overall ratio of theoretical versus empirical explanations would average out the differences in response times according to solution strategy. Nevertheless, the results obtained do have some implications for the assessment of reasoning performance of bilinguals.

The fact that answers that received a theoretical justification had higher response times does indicate that if theoretical explanations are given, reasoning takes place. It also appears that subjects in higher grades give theoretically justified answers more quickly than those in the lower grades, reflecting an increasingly routinized application of reasoning skills as children get older. Response times for empirically grounded answers, on the other hand, stay the same, leading to higher response times for empirically grounded answers than for the theoretical ones in the fifth grade. The mental operations

underlying empirical justifications, thus, appear to differ according to grade level: In the lower grades, no reasoning underlies empirically justified answers, whereas in the higher grades, subjects do something in addition to the routine-like application of reasoning skills. It is not inconceivable that subjects in the higher grades try to solve problems using just their reasoning skills, and use empirical strategies as a default option if reasoning does not lead them to an answer.

This impression receives some support if one considers the distribution of response times according to grade for the answers for which no explanation at all is given (the 'don't know' category). As in the case of the empirically justified answers, response times go up for this type of explanation (i.e. no explanation) as grade gets higher. They are higher in any case than the response times for empirically grounded answers reflecting an extra effort in all grades. In the third and fourth grades, they are lower than the theoretically grounded answers. As subjects reason, then, they do so to a lesser extent. In the fifth grade, however, response times are dramatically higher in the no explanation case, a tendency which, again suggests that fifth graders consider alternative options to meet the task demands if reasoning does not work. This tendency reflects an understanding in the fifth grade that reasoning is the main concern in the task. The results for the sixth grades show a reversal of the tendency noted above. They show that, while

response times for the theoretically justified answers are longer, empirically justified answers, and answers not justified are given more quickly. It is likely that the weaker proficiency in English made it more difficult for this group to apply reasoning skills in a routinized fashion, and that response times for theoretically justified answers go up as a result. The lower response times for the other justification categories are probably due to the reduced effort in the weaker language, noted above, which should be more pronounced for this group if the results obtained are indeed attributable to their weaker language proficiency.

That subjects do come up with empirical explanations in spite of this understanding may be a result of the inconsistency perceived by subjects in the higher grades between (dis)confirming a conclusion and not being able to indicate why they do so, reflecting a more general awareness of the task demands. In either case, the pattern of response times according to grade level and explanation type indicates an increasing awareness of what the task is all about as grade gets higher, and it indicates an increasing amount of ease in the use of reasoning skills to solve the task. Whether empirical justifications reflect the absence of reasoning appears to depend on grade. In the lower grades, empirical justifications seem to detect limited task understanding whereas in the fifth grade, they seem to be attributable to a failed reasoning effort.

The difference between reasoning skills and task understanding appears to be gradual rather than absolute. Most children display reasoning skills to a more or lesser extent, and an increased awareness of the task demands seems to lead children in the higher grades to consider reasoning strategies first in order to solve the syllogisms, and to consider alternative strategies if reasoning does not lead them anywhere. For the earlier grades, this does not appear to be the case, although even in the earlier grades, children may intuitively feel what the task requirements are without being able to meet them using the appropriate solution strategies. The impression that task understanding is present in the earlier grades in some rudimentary form is consistent with Hawkins et al.'s (1984) finding that young children do show an understanding of the purpose of the task, if only under specific conditions. The conditions in Hawkins et al.'s study were the administration of 'fantasy' syllogisms, syllogisms which were clearly unrealistic, before administering syllogisms that could be related to everyday experience. For younger children to understand a syllogism task, thus, it seems to be required that the counterfactual nature of the task is made clear to them by letting them solve a number of syllogisms in which they are prevented from applying their factual knowledge. The necessity of understanding the counterfactual nature of the task in order to meet its demands to reason (see also Bloom, 1981) implies that more reasoning ability

may have been displayed by the younger subjects in this sample if a fantasy condition had been included. In other words, it is not unlikely in light of Hawkins et al.' findings, that reasoning performance in the lower grades would have been better in this study if the testing condition had facilitated an understanding of the counterfactual nature of the task.

In light of these considerations, it is not surprising that there is no joint contribution of solution accuracy and explanation type to the variation in response times. The predictions of response times by solution accuracy would be different at different grade levels due to the differences in mental operations that underlie the two explanation types. As a consequence, variations according to those factors have probably been evened out due to the opposing tendencies in response times according to explanation type.

Although response times do not differ significantly according to solution accuracy, they do differ if language of administration is taken into consideration. In the English administration, there is hardly any difference in response times between right and wrong answers, whereas there is a difference in Spanish. Having found no interaction between language of administration and solution strategy, this difference does not imply that more reasoning takes place in Spanish. Instead, it seems more likely that the 'default strategy' is more frequently

used in Spanish than in English: In Spanish, subjects are more likely to consider other strategies to solve the syllogisms if reasoning does not lead them anywhere, suggesting that subjects are more thorough in their stronger language, and more flexible in their approach to the task. The difference according to grade in the two languages is likely to be attributable to the difference in proficiency in the two languages. Proficiency measures have not been included in the analysis, but it appears that fourth graders are relatively weak in English, and that fifth graders are relatively strong in English, and relatively weak in Spanish. Whatever the case may be, there clearly is an opposing tendency in the response times upon comparison of English and Spanish, a tendency which suggests that the more thorough people are in their stronger language, the less so they are in their weaker one, regardless of whether they reason or not.

Having concluded that solution accuracy is not directly affected by explanation type, we are left to explain why subjects performed above chance level regardless of the factors taken into consideration in this analysis. The most likely explanation appears to be that the syllogism characteristics themselves account for this result. All syllogisms that contained a denial led to a negative conclusion, and all syllogisms that did not contain a denial led to an affirmative conclusion. Having no denial necessarily leads to an affirmative conclusion for any

syllogism, unless it is invalid, which none of the syllogisms were in this study. Although syllogisms that do contain a denial do not necessarily get a negative answer (e.g. one can propose 'some A are not B', and have an affirmative conclusion), most syllogisms with a denial in the premises do have a negative conclusion (and in my analysis, all of them had). This phenomenon partly constitutes what has been referred to in the literature as the 'atmosphere effect' (Woodworth & Sells, 1935).

Before accepting this effect as an explanation for the inflated accuracy scores for our subjects, an alternative, and more specific explanation should be considered, namely that the same artifact is attributable to processing errors, a possibility that has been suggested by Chapman and Chapman (1959). In the latter explanation, a false positive result for syllogistic reasoning is obtained due to the fact that during processing of the syllogism, content terms within one premise have been illicitly converted. For example, if a premise reads 'All A are B', the premise would be retained for reasoning as 'All B are A'. Such conversion does not reduce the likelihood of giving the right answer for most syllogisms.

There is an important difference, however, between the atmosphere effect and illicit conversion: In the latter case, it is assumed that reasoning takes place, and that the syllogism is actually processed as such, whereas the occurrence of the

atmosphere effect does not imply this. In the present study, reasoning does take place in some cases but not in others, and consequently, we can conclude that there is evidence for the atmosphere effect (performance above chance level without reasoning), and not enough evidence for illicit conversion (performance above chance level in spite of reasoning mistakes), since the present design does not enable us to detect mistakes in reasoning.

In order to get a better grip on the reasoning process itself, it would be necessary in subsequent work to make a distinction between instances of reasoning and instances of the use of other skills to solve reasoning tasks, and make a separate assessment of the subjects' level of performance. A good way to do this would be to enable subjects to draw their own conclusions (Johnson-Laird, 1983) rather than having them confirm or disconfirm a given one. Moreover, including invalid syllogisms may provide additional information about the level of reasoning proficiency of the subjects (Orasanu & Scribner, 1982). Furthermore, to reduce the effect of memory load on reasoning performance, it is worth replicating the current study using a paper and pencil format rather than an oral administration. The more general implication of the results reported above is that an assessment of what children can do, does not merely depend on the rightness or the wrongness of their performance, especially if this performance is not in the subjects' native language.

Regardless of the artifacts noted above, and the fact that using binary answering categories leads to a high chance level (50%), there is a contradiction in the results if solution accuracy and type of solution strategy do not affect each other whereas only one type of strategy should work. As the present results indicate, using the right strategy does not lead to more right answers, and consequently, rightness or wrongness in task performance does not automatically imply the possession of a certain skill. This contradictory result is worth keeping in mind both in cognitive research and in classroom situations. For cognitive research, it indicates that to assess the possession of a certain skill, an assessment of the appropriateness of problem solving strategies is likely to be more effective than an assessment of task success. Moreover, it appears that the strategies children use for problem solving can manifest themselves in different ways due to developmental factors. The implication of the present study for classroom situations is essentially the same: It is not enough to know whether students do things right, unless we also know whether they are doing them right in the right way.

References

Bloom, A. (1981). The Linguistic Shaping of Thought: A Study in the Impact of Language on Thinking in China and the West. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chapman, L. J. & Chapman, J. P. (1959). Atmosphere effect re-examined, Journal of Experimental Psychology, 58, 220-226.

Cole, M., Gay, J., Glick, J. & Sharp, D. (1971). The Cultural Context of Learing and Thinking. New York: Basic Books.

Cole, M. & Scribner, S. (1981). The Psychology of Literacy. Cambridge, MA: Harvard University Press.

Hawkins, J., Pea, R. D., Glick, J. & Scribner, S. (1984). "Merds that don't laugh like mushrooms": Evidence for deductive reasoning by pre-schoolers, Developmental Psychology, 20, 584-594.

Johnson-Laird, P. N. (1983). Mental Models. Cambridge, MA: Harvard University Press.

Koopmans, M. (in preparation). Reasoning in two languages: An assessment of the effect of language proficiency on the reasoning performance of Puerto Rican bilingual elementary

school children. Doctoral Dissertation, Harvard University.

Luria, A. L. (1976). Cognitive Development: It's Cultural and Social Foundations. Cambridge, MA: Harvard University Press.

Orasanu, J. & Scribner, S. (1982). The development of verbal reasoning: Pragmatic, schematic, and operational aspects. In W. Frawley (Ed.) Linguistics and Literacy. New York: Plenum.

Piaget. J. (1967). Six Psychological Studies. New York: Random House.

Scribner, S. (1977). Modes of thinking and ways of speaking: Culture and logic reconsidered. In P. N. Johnson-Laird & P. C. Wason (Eds.) Thinking: Readings in Cognitive Science. Cambridge: Cambridge University Press.

Sternberg, R. J. (1986). Toward a unified theory of human reasoning, Intelligence, 10, 281-314.

Woodworth, R. S. & Sells, S. B. (1935). An atmosphere effect in formal syllogistic reasoning, Journal of Experimental Psychology, 18, 451-460.